

ABSTRACT

Systems and methods that facilitate dimensional transformations of data points are disclosed. In particular, the subject invention provides for a system and methodology that simplifies dimensional transformations while mitigating variations of a distance property between pairs of points. A set of n data points in d dimensional space is represented as an $n \times d$ input matrix, where d also corresponds to the number of attributes per data point. A transformed matrix represents the n data points in a lower dimensionality k after being mapped. The transformed matrix is an $n \times k$ matrix, where k is the number of attributes per data point and is less than d . The transformed matrix is obtained by multiplying the input matrix by a suitable projection matrix. The projection matrix is generated by randomly populating the entries of the matrix with binary or ternary values according to a probability distribution. Unlike previous methods, the projection matrix is formed without obtaining an independent sample from a Gaussian distribution for each entry in the projection matrix, without applying a linear algebraic technique to generate the projection matrix and without employing arbitrary floating point numbers. Processes and/or algorithms can utilize the reduced transformed matrix instead of the larger input matrix to facilitate computational efficiency and data compression.